

```

1   TTCTCTCTGT TTGCTTACTC CCTATCCGGG GGGCCAAGGC GCTGTCTCCG
51  CCGCCCAAGC CCGCGTAAA CCTGGGTGAC CTCGGAGACA TCCGTTGGAG
101 CATGAGTTCC CGACATCAGG CGGCGGCGGT GGTCCGGGAG AAACCCGGCG
151 GCGGGGAGAT AAGCCTGCCC AGGAGGCAGG GGGCTGGGCT AGCTGCCCCG
201 CCGCGGCGCT GACTTCGTTG GGGAGGGAGA CGCCCGGCTC CCGCCCTAA
251 CTAGCCAGC CGCGCGGAGC GCCTGGGAGA GGAGAAGGAG CCGACCTGCC
301 GAGATGGAGG CGACCGGCAC CTGGGCGCTG CTGCTGGGCG TGGCGCTGCT
351 CCTGTGCTG ACCTGGGCGC TGTCGGGGAC CAGGGCCCGA GGCCACCTGC
401 CCGCGGGGCC CAGCGCGCTA CCACTGCTGG GAAACCTCCT GCAGCTACGG
451 CCGGGGGCGC TGTATTACAG GCTCATGCGG CTGAGTAAGA AGTACGGACC
501 GGTGTTCACC ATCTACCTGG GACCTGGCG GCCTGTGGTG GTCCTGGTTG
551 GGCAGGAGGC TGTGCGGGAG GCCCTGGGAG GTCAGGCTGA GGAGTTCAGC
601 GCGCGGGGAA CCGTAGCGAT GCTGGAAGGG ACTTTTGATG GCCATGGGGT
651 TTTCTTCTCC AACGGGGAGC GGTGGAGGCA GCTGAGGAAG TTTACCATGC
701 TTGCTTGTCC GACCTGGGC ATGGGGAAGC GAGAAGGCGA GGAGCTGATC
751 CAGGCGGAGG CCGGTGTCT GGTGGAGACA TTCCAGGGGA CAGAAGGACG
801 CCCATTCGAT CCTCCCTGC TGCTGGCCCA GGCCACCTCC AACGTAGTCT
851 GCTCCCTCCT CTTTGGCCTC CGCTTCTCCT ATGAGGATAA GGAGTTCAG
901 GGTGTGGTCC GGGCAGCTGG TGGTACCCTG CTGGGAGTCA GCTCCAGGG
951 GGGTCAGACC TACGAGATGT TCTCCTGGTT CCTGCGGCCC CTGCCAGGCC
1001 CCCACAAGCA GTCCTCCAC CACGTCAGCA CCTTGGCTGC CTTACAGTC
1051 CCGCAGGTGC AGCAGACCA GGGGAACCTG GATGCTTCGG GCGCCGACG
1101 TGACCTTGTG GATGCCCTCC TGCTGAAGAT GGCACAGGAG GAACAAAACC
1151 CAGGCACAGA ATTCACCAAC AAGAACATGC TGATGACAGT CATTTATTTG
1201 CTGTTTGCTG GGACGATGAC GGTGAGCACC ACGGTGCGCT ATACCTCCT
1251 GCTCCTGATG AAATACCCTC ATGTCCAAA GTGGGTACGT GAGGAGCTGA
1301 ATCGCGAGCT GGGGGCTGGC CAGGCACCAA GCCTAGGGGA CCGTACCCGC
1351 CTCCCTTACA CCGACGCGGT TCTGCATGAG GCGCAGCGGC TGCTGGCGCT
1401 GGTGCCCATG GGAATACCCC GCACCTCAT GCGGACCAAC CGCTCCGAG
1451 GGTACACCCT GCGCCAGGGC ACGGAGGTCT TCCCCCTCCT TGGCTCCATC
1501 CTGCATGACC CCAACATCTT CAAGCACCCA GAAGAGTTCA ACCCAGACCG
1551 TTTCTGGAT GCAGATGGAC GGTTCAGGAA GCATGAGGCG TTCCTGCCCT
1601 TCTCCTTAGG GAAGCGTGC TGCTTGGAG AGGGCTGGC AAAAGCGGAG
1651 CTCTTCTCT TCTTACCAC CATCCTACAA GCCTTCTCCC TGGAGAGCCC
1701 CTGCGCGCGG GACACCCTGA GCCTCAAGCC ACCGTCAGT GGCCTTTTCA
1751 ACATTCCCC AGCCTTCCAG CTGCAAGTCC GTCCCACTGA CCTTCACTCC
1801 ACCACGCAGA CCAGATGAA GAAGGCAACT TGAAGTGGT GGGTGCCAG
1851 GACGCTGCCT CCAGCTCAA CAGTGGGCAT GGACAGGGT AATGTCTCCA
1901 GAGTGTACAC TGCAGGCAGC CACATTTACA CGCTGTCAGT TGTTTTCCGG
1951 AGTCTGTCCC ACGGCCACA CGCTCACTG ACTCATGCTG CTAAGATGCA
2001 CAACCGCACA CCATACACA ACTACAAGG CCACAAAGCA ACTGCTGGGT
2051 TAGCTTTCCA CAGACATAA TATAGTCCAT CTGCAATCAC AAGCACATAG
2101 CCAGGTAACC CACCAACTCC CCTGGATCTG CAGCCACAC GTGGGAGTCT
2151 GGCTGTACAC TTCACAAGCC ACAGAAACGG CCACACATGT TCACAGCTCA
2201 CACGCCCTCT CCATTATCG AACTTCTCAG TGTCCCTGTC CCTGGTGCCT
2251 GGCACAGGGA ACAGCATGCC CCTCCGGGG TCATGCCACC CAGAGACTGT
2301 CGCTGTCTAT GGGCCCAACT CATGCTCCCT CTCTTGGCTA CACCACTCTC
2351 CCAGCTGTG ACCACCGATG TCCACACACC CCAACCACT TGTCCACACA
2401 GCTACCCACG TACGACATCG TCCTGGCTCC CCAGAGTATC TTCCCACTGA
2451 GACACGCGCG CCCACAGAG GCACAGTCCC CAGCCACCTC TGCAACTGCA
2501 GCCCTCAGTC ACCCCTTTTT AAGCACCTG ATTCTACCAA ATGCAAAAC
2551 ATCTGGGTCT GCATTATGC ACAGAGACTT TGGACATACG AGGACCTCA
2601 GACCGGAGGA ACACCTGCC AACCCCAACA CGTGCTTATG TAACCACGTG
2651 GAAAGCGGCC CCTGCTGCC CTCCACACAC ACATACACAC TCACTGATCT
2701 ACAGCCCCTG TTCGGCGTCA GAGTCCCCAC TAGACCCAGT GGAAGGGGT
2751 AGAGACCAAG TAGGGGCCAG TTTCCAATTC ACCCTGTGAG GGAGTGAGCC
2801 GGATCTGACG TTCCTTGTGA CTTAAGGGTC CGGCTTGGGA ATTAAGTTT
2851 GTTTCTGGCC TTTAGCCTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
2901 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA

```

**FEATURES:**

5'UTR: 1-303  
Start Codon:304  
Stop Codon: 1815  
3'UTR: 1818

# HOMOLOGOUS PROTEINS:

gi 117254 sp P24461 CPG1_RABIT CYTOCHROME P450 2G1 (CYPIIG1) (P...	516	e-145
gi 404777 gb AAA31432.1  (L10912) cytochrome P-450 2B-Bx [Oryct...	513	e-144
gi 479930 pir  S35666 cytochrome P450 2B4 isoform Bx - rabbit >...	513	e-144
gi 117212 sp P00178 CPB4_RABIT CYTOCHROME P450 2B4 (CYPIIB4) (P...	508	e-143
gi 320075 pir  S31277 cytochrome P450 2B4-B1 - rabbit >gi 21369...	508	e-143
gi 4731350 gb AAD28466.1  (AF128849) cytochrome P450 2B10 relat...	508	e-143
gi 3123191 sp P04167 CPB2_RAT CYTOCHROME P450 2B2 (CYPIIB2) (P4...	508	e-143
gi 117213 sp P12789 CPB5_RABIT CYTOCHROME P450 2B5 (CYPIIB5) (P...	507	e-142
gi 89973 pir  A27717 cytochrome P450 2B5, hepatic (form HP1) - ...	507	e-142
gi 2144292 pir  O4RTP2 cytochrome P450 2B2 - rat	506	e-142

## EST:

gb BE148597 BE148597 MR0-HT0241-150500-010-b02 HT0241 Homo sapi...	1091	0.0
gb BF359243 BF359243 RC6-ET0072-150600-011-F01 ET0072 Homo sapi...	779	0.0
gb AW753778 AW753778 RC1-CT0286-301099-011-f08 CT0286 Homo sapi...	652	0.0

## EXPRESSION INFORMATION FOR MODULATORY USE:

gb|BE148597| head\_neck  
 gb|BF359243| lung\_tumor  
 gb|AW753778| colon

## Tissue expression:

Human leukocyte

1 MEATGTWALL LALALLLLLT LALSGTRARG HLPGPPTPLP LLGNLLQLRP  
51 GALYSGLMRL SKKYGPVFTI YLGPWRPVVV LVGQEAAREA LGGQAEFSG  
101 RGTVMLEGT FDGHGVFFSN GERWRQLRKF TMLALRDLGM GKREGEELIQ  
151 AEARCLIVET QGTEGRPFDP SLLLAQATSN VVCSLLFGLR FSYEDKEFQA  
201 VVRAAGGTLL GVSSQGGQTY EMFSWFLRPL PGPHKQLLHH VSTLAAFTVR  
251 VQVQHQNLD ASGPARDLVD AFLKMAQEE QNPGTEFTNK NMLMTVIYLL  
301 FAGTMTVSTT VGYTLILLMK YPHVQKWVRE ELNRELGAGQ APSLGDRTRL  
351 PYTDAVLHEA QRLALVPMG IPRTLMRTTR FRGYTLPGT EVFPLLGSIL  
401 HDPNIFKHPE EFNPDRLDA DGRFRKHEAF LPFSLGKRVC LGEGLAKAEL  
451 FLFFTTILQA FSLESPCPPD TSLKPTVSG LFNIPPAFQL QVRPTDLHST  
501 TQTR

# FEATURES:

## Functional domains and key regions:

Prosite search results:

[1] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE

CAMP- and cGMP-dependent protein kinase phosphorylation site

128-131 RKFT

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE

Protein kinase C phosphorylation site

Number of matches: 6

1 61-63 SKK  
2 99-101 SGR  
3 248-250 TVR  
4 288-290 TNK  
5 378-380 TTR  
6 473-475 SLK

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE

Casein kinase II phosphorylation site

Number of matches: 3

1 119-122 SNGE  
2 192-195 SYED  
3 343-346 SLGD

[4] PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 10

1 51-56 GALYSG  
2 109-114 GTFDGH  
3 115-120 GVFFSN  
4 188-193 GLRFSY  
5 207-212 GTLLGV  
6 257-262 GNLDAS  
7 284-289 GTEFTN  
8 339-344 GQAPSL  
9 370-375 GIPRTL  
10 444-449 GLAKAE

[5] PDOC00009 PS00009 AMIDATION

Amidation site

Number of matches: 2

1 140-143 MGKR  
2 435-438 LGKR

[6] PDOC00029 PS00029 LEUCINE\_ZIPPER

Leucine zipper pattern

Number of matches: 2

1 32-53 LPPGPPTPLLLGNLLQLRPGAL  
2 39-60 LPLLGNLLQLRPGALYSGLMRL



# BLAST Alignment to Top Hit:

```
>gi|117254|sp|P24461|CPG1_RABIT CYTOCHROME P450 2G1 (CYPIIG1)
(P450-NMB) (OLFACTIVE)
pir||B31944 cytochrome P450 2G1 - rabbit
Length = 494
```

Score = 516 bits (1315), Expect = e-145  
Identities = 248/491 (50%), Positives = 345/491 (69%), Gaps = 3/491 (0%)

```
Query: 1 MEATGTWALLLALAL-LLLLTLALSGTRARGHLPPGPPTPLPLLGNLLQLRPGALYSGLMR 59
ME G + + LAL LL+ +A + G LPPGPPT+P LGNLLQ+R A + ++
Sbjct: 1 MELGGAFTIFLALCFSCLLILIAWKRVQKPGRLPPGPPTPIPFLGNLLQVRTDATFQSFLK 60

Query: 60 LSKKYGPVFTIYLGFWRPVVVLVGQEAAREALGGQAEFSGRGTVAMLEGTFDGHGVFFS 119
L +KYGPVFT+Y+GP RPVV+L G EAV+EAL +A+EFSGRG +A +E F GHGV +
Sbjct: 61 LREKYGPVFTVYMGP-RPVVILCGHEAVKEALVDRADFSGRGELASVERNFQGHGVALA 119

Query: 120 NGERWRQLRKFTMLALRDLGMGKREGEELIQAEARCLVETFQGTGEPFDPSSLQAQATS 179
NGERWR LR+F++ LRD GMGKR EE IQ EA L+E F+ T+G P DP+ L++ S
Sbjct: 120 NGERWRILRRFSLTILRDFGMGKRSEIERIQEEAGYLLEEFKTKGAPIDPTFFLSRTVS 179

Query: 180 NVVCSLLFGLRFSYEDKEFQAVVRAAGGTLGVSQGGQTYEMFSWFLRPLPGPHKQLLH 239
NV+ S++FG RF YEDK+F +++R + + +S+ Q Y+M+S ++ LPG H ++ +
Sbjct: 180 NVISSVVFGSRFDYEDKQFLSLLRMINESFIEMSTPWAQLYDMYSGVMQYLPGRHNRIYY 239

Query: 240 HVSTLAAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYL 299
+ L F +V+ ++ +LD P RD +D FL+KM Q++ NP TEF KN+++T + L
Sbjct: 240 LIEELKDFIAARVKVNEASLDPQNP-RDFIDCFLIKMHQDKNNPHTEFNKLNVLTLNL 298

Query: 300 LFAQTMTVSTTVGYTLLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDAVLHE 359
FAGT TVS+T+ Y LL+MK+P VQ + EE+N+ +G + PS+ DR ++P+TDAV+HE
Sbjct: 299 FFAGTETVSSTLRYGFLIMKHPEVQTKIYEEINQVIGPHRIPSVDDEVKMPFTDAVIHE 358

Query: 360 AQRLALVPMGIPRTLMTTRFRGYTLPGQTEVFLLGSILHDPNIFKHPEEFNPDRFLD 419
QRL +VPMG+P ++R T FRGY LP+GT+VFPLLS+L DP F HP++F P FLD
Sbjct: 359 IQRLTDIVPMGVPHNVIRDTHFRGYLLPKGTDVFLLGSVLKDPKYFCHPDDFYQHFELD 418

Query: 420 ADGRFRKHEAFLPFSGLKRVCLGEGLAKEAELFFFTTILQAFSLESPPDLSLKPTVS 479
GRF+K+EAF+PFS GKR+CLGE +A+ ELFL+FT+ILQ FSL PP + + P +S
Sbjct: 419 EQGRFKKNEAFVPFSSGKRICLGEAMARMELFLYFTSILQNFSLHPLVPPVNIIDITPKIS 478

Query: 480 GLFNIPPAFQL 490
G NIPP ++L
Sbjct: 479 GFGNIPPTYEL 489
```

## Hammer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00067	Cytochrome P450	594.4	6.9e-175	1

## Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00067	1/1	33	493	1	497	594.4	6.9e-175

1	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
51	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
1901	NNNNNNNNNN	NNNNNNNNNN	NNNNNTGACA	GGGGCCATGA	TGGAGACACC
1951	TTGGATCGAA	GAGGTCACAG	CACCCTCCTC	TTTCTTCCTC	CCTACCCCCA
2001	GCTGAGTAAG	AAGTACGGAC	CGGTGTTTAC	CATCTACCTG	GGACCTTGCG
2051	GGCCTGTGTT	GGTCTGGTT	GGGCAGGAGG	CTGTGCGGGA	GGCCCTGGGA
2101	GCTCAGGCTG	AGGAGTTCAG	CGGCCGGGGA	ACCGTAGCGA	TGCTGGAAGG
2151	GACTTTTGAT	GGCCATGGTA	AGTCAAGGGC	TGCTAGGCCC	TCCGCTCACA
2201	GCCTGCCACC	ACTTACTGGT	GTGTGACCTT	TGCACATGGC	TTAGTCCCTC
2251	TGTTGCCTCA	TCTGTCAAAT	GGAGTGATAA	CAGTGCCCAT	GACCCGGGTG
2301	CAGTGGCTAG	TGCTTGAAAT	CCCAACACTT	TGGGAGGCGG	AGGTGGGTGG
2351	ATCACTTGAG	GTGAGGAGTT	CGAGACCAGC	CTGGCCAACA	TGGTGAAACC
2401	CTGTCTCTAC	TAAAAATATA	AAAATTAGCT	GGGCATGGTG	GTGCGTACCT
2451	GTAAATCCCA	ATACTTGGA	GTTGAGGCA	GGAGAATCGC	TTGAACCCGG
2501	GAGGCAGATG	TTGCAGTGAA	CCAAGACTGT	GCCACTGCAC	TCCAGTCTGG
2551	GCAACAGAGT	GAGCCTCCAT	CTCAAACAAA	CAAACAAAAA	GCAGTGCCCA
2601	TCATGTAGGA	TTGAGTGATT	GAGTGAGGAC	TGAGCCTTGT	GCAAAGTGAG
2651	CACTCACTAA	TCACCAGGTT	GATGATCAG	TGATAACCAT	CAATGATCCA
2701	GGTAAAGCCC	TGAGGGTTCA	GAAAGATGCC	GGAGCGCTTT	CAAGGTGCTG
2751	GGGATTGGTG	GGCAAGCCCT	CGAATAATAG	AAACAGTTCT	CTGTATTACA
2801	ACAGAAAGCA	GGAGGCCCAT	GCTGGGTGCT	GCCAGGAAT	CAGTAGTAAC
2851	TAAGACAGCA	CCGGTGCTGC	TTCCCCAGCG	CACCTAGGCC	AGTGGGGAAA
2901	CAGACTCACC	ACACAGTCCC	AGCCCAGAGT	GGTCAGGGCC	AAGATGGGGA
2951	AGCACGGGGG	GAAAGGTCAG	GGTGGGATGG	GGAGGGGTCA	GGGCAAGAGG
3001	GGTCAGGGCG	AGGCTGAGGG	AAGCCCTGGG	ACTGTAGGAA	TTTAGAGGAG
3051	GTACCTGACC	CGGCATGTTT	GGTGAGGGAG	ATTCAGGAAG	TCTTCTTGGA
3101	AGAGAGGCTG	TCGGAGCTGA	GACTCATAAG	ATGAGTGGGG	AGGGTGTTC
3151	AGGCAGAAAG	ACCAGCACCT	ACAAAAGCAT	GACTTTGAGA	GAAGCATTCA
3201	TCCATTCAAC	TGATGAATTT	TCAGACTGGG	CACGCTGGCT	CATGCCTGTA
3251	ATCCGAGCAC	TTTGGAAAGC	TGAATGGGGA	GGATGACTTG	AGCCTAGGCA
3301	TTTGTGACAA	GCCTGGGCAA	CATGGTGAGA	CCCTGCCTCC	ACAAAACAAA
3351	CAAACAAACA	AAAAATCATT	ATACCTGGTA	CCATGGGTAC	CAGGTACATA
3401	GAAATGACTC	AGGCAGATAT	GGTGTCTCT	CCTACTGTGC	GAGAGGCGGG
3451	CTTATACTGC	AGTAAGACAA	TAGAGGGAGG	GAATATAATC	CTAAAATGAG
3501	AGGTACAGAT	TTGAGAGCAA	ACACAGGGCA	CAGGCATATG	TACGAGGGTA

FIGURE 3, page 1 of 10

```

3551 AAGAGGGAAT CAGGGAAGGC TTCTCAGAGA AGGTGACATT TAAGCCGGGA
3601 CATGAAGGAT GAACGAGTTA GTTCACCAAG GATGGGATGG AAAGGGGTGA
3651 GAGTGATGGA GGCAGAGGGA ACTGCAGGAT CATAGGCCTA GACAGGGGAT
3701 CCTGACGCCC TTGAGGAAGT GAGAGAAGAC CAGCGCAGTC GTAGTGGGTT
3751 AAGTAACAAA GCTGAGAAGC CAGGGAATC CCTGGTCATG CAGGGCCTGT
3801 GAGTCACGTC AGAGTGTGTT GGCTTTTGTT TTTCTCTGGG GCAGTCGATT
3851 TTAAGCAGGG AACAGCTGTA TTCAGAGTTG GGAAGATCCT GTGGTTGCTG
3901 CCTGAAGGGG ATGAAACTGG AGGCTAGGAG CCCAGGGTGA TAGGGAGGAT
3951 CCAGGGTGAT GGGGAGGCTG GGAGGTCCGC GGTGATGGAC CAGGGCTGGG
4001 GCCAGGGGAT GGGGAGGAAG GAGTAATTGG GAGAGGCCTG GGGCTCTGGC
4051 CGAGGAATGG ATGGTGGGCT GAAACAGGGA GAGGAGAGAT GCTTAGGCCA
4101 CTTTGGAAAC CAGTAGGGCA AGGACAGGAG ACACCCAAGG GGAAGTGCCC
4151 AAGAGACCAC GACAGGCTGG CATTGGACAG GGAAGGTCTG TCTGGAGCAG
4201 GTGTCTTGGA TAAGGGAGGA AAATGGTGCA GTTCCATCCT CCTCCCTCTC
4251 TGTTCAACCT CTAAACTACA TGGGGCACAG GACCCAGTGG GACTCCATAA
4301 ATGATGGGAT GGGTGGATGG AAGGAAGGAA GGAGGAAACA ACTCTTCAT
4351 CATCTGGTT ATTTACAGAA CAGGCCAGGT GCGGTGCTCA CGCTTGCCAT
4401 TCTAGCACAT TGGGAGGCTG AGGTGGGTGG ATTACCTCAG GTCAGGAGTT
4451 CAAGACCACG CTAGACAACG TAGAGAAACC CCATCTCTAC TGAAGATATA
4501 AAATTAGCTG GCGTAGTGG CATATGCCTG TAATCCAGC TAGTCGGGAA
4551 GCTGAGGCAG GAGAATCGCT TGAACCCGAG AGGCAGAGGT TCGCGTGAGC
4601 TGAGATCGTG CCATTGCACT CCAGCCTGGG TGACAAAGCA AGACCTCGTC
4651 TCAATAATAA TAATAATTAC AAAACAGAAG GAGCCTGGGT CATCCAGCT
4701 ACCTACTTTT CAGGAGAATG TACTCCCTTA CCCAAGGGCA AAGGATGGGA
4751 GAACCAAGTT GATTATGCAT TTATTGAGCA CTTACTGAGT CCTCATCCCT
4801 GGGCTAGGCT GGAATGGACT CAGATGGAGC CTGAAGAGTC CCCCTCAGGG
4851 AACCGGGTGA GAAAGAAGGA GGAATCGGCC GGGCGCGGTG GCTCACGCCT
4901 GTAATCCCAA CACTTTGGGA GGCTGAGGTG GGTGGATCAC AAGGTCAGGA
4951 GATCGAGACC ATCCTGGCTA ACACAGTGAA ACCCCATCTC TACTAAAAAT
5001 ACAAATAATG AGCCAGGCAT GGTGGCGGGC GCCTGTAGTC CCAGCTACTC
5051 AGGAGGGTGA GGCAGGAGAA TTGCTTGAAC CCGGGAGGCA GAGGTTGCAG
5101 TGAGACGAGA TCACGCCACT GCACTCCAGC CTGGGCAACA GAGCGAGATT
5151 CCGTCTCAAA AAAAAAAGA AAGAAAGGAA GAAGGGGGAA TGGGGGAGAG
5201 GGGCCGGTCC CTTTTGAGT CTAGCCTTCT GCGCAGGGGT TTTCTTCTCC
5251 AACGGGGTGA GGTGGAGGCA GCTGAGGAAG TTTACCATGC TTGCTCTGCG
5301 GGACCTGGGC ATGGGGAAGC GAGAAGGCGA GGAGCTGATC CAGGCGGAGG
5351 CCCGTGTGCT GGTGGAGACA TTCCAGGGGA CAGAAGGTCA GCATGGCGGG
5401 GTCACCCAGG GGTCTCCAGC CGAGTGAAG GGAAACTCT CCTACTGTGG
5451 CTGGGGGTGA CCCCACCCA GGTCTGGAA TGGGCAGGAG GGAAGCCTT
5501 GAACTCTAGG GCTGGCCTGG GGGTCTGTT CACTGCCACC TTCTGTCTCT
5551 GTCCCACTGT CTCTCCGAGG CTGTCTATGAC ATCTCTCTGT GTGTCTCTGG
5601 TGCTATCATC CCATTCTTCC TGGGTCTCCA TCTCTCTCTC TGCTCTCTTT
5651 CTTTCTCTCT CTTTCTCTCT ATTTTGTGGG CCCTCAGTCT ATCTCTGTTT
5701 CTGTCTCCCT GTCTGTGTGA TGGTCACTCT GTTCTTTCT CCCTGTCTGT
5751 TTCTCTGTCC CTATCTGTCT GTATCCTTCT TTGCCTGTTT AGCTCTCTCC
5801 CTGCGCTGTC CATCCATCTT TCCCTGCCTC CTTGTCTCTC TCTGGTTGGG
5851 TTCAGCCCCA ACCTGCTCCC CTCTGCTGG CTCCATCACA GCCTACCTCC
5901 CTGCCCCCAT TCCCCCAGG ACGCCCATTC GATCCCTCCC TGCTGCTGGC
5951 CCAGGCCACC TCCAACGTAG TCTGCTCCCT CCTCTTTGGC CTCCGCTTCT
6001 CCTATGAGGA TAAGGAGTTC CAGGCCGTGG TCCGGGCAGC TGGTGGTACC
6051 CTGCTGGGAG TCAGCTCCCA GGGGGGTGAG GTGAGTGGGT GGGACCCCTC
6101 TCCAACCTACC TTCCCTGAAG GTTCCTGCCA AGGTCCCATG AGAACTAGCT
6151 GCCCTTCTCC CCACAGACCT ACGAGATGTT CTCTGTGTTT CTGCGGCCCC
6201 TGCCAGGCCC CCACAAGCAG CTCCTCCACC ACGTCAGCAC CTTGGCTGCC
6251 TTCACAGTCC GGCAGGTGCA GCAGACCAG GGAACCTGG ATGCTTCGGG
6301 CCCCCACGTC GACCTTGTCG ATGCCTTCCT GCTGAAGATG GCACAGGTGT
6351 GGGAAAGGGT CAGGGACCCC CTCTCTGAAT GGGCGTGGTG ACCTGGCAGG
6401 TCCCAGCCAG GTGTCCCTGG GGACCTCAAT TGGGTTCCCT TCTCTTTCTC
6451 TCTCTGCATG TCTCTGTGAG TATGAGTGTC TCTGTGCATG TGTGTGCATC
6501 CCTTCTCTGC ACATCTGTGC TGGCCCTTTC AGGGCGTGGC TCTCACTGCC
6551 TCTCCCGCCC CCGACCTGGG CATTGTGTC GGGCTGTCTG TCTCTCCAGC
6601 ATCTCTCTCT TTTCTCCCTC CCACCTCGGC CCTGTGTTT AGGCCCCATG
6651 CCCAGGGTCC TACACCAGCA ATCCCAGGA TCACTTCATC CCATCCCCTG
6701 CAGCCTCCCC AGACTTTTAT GTAAATTCAC AATTTTATGT GAATATAGGT
6751 CATTTATTAG GAAGCCTTGC AATATCAAGT TATGTTAATA AAGTCCACTT
6801 TATTAATTAT ATAAGAACAA TATTTCTTTT CCTTTTTTTT TTCTTTTCTT
6851 TTTAAAGAGA CAGGATCTCT TTCTGTTGCC CAGGCTAGAG TACAGTTGCA
6901 AAATCATAGC TCACTGCAAC CTTGAACTCC TGGGCTCAAG CAATCCTCCT
6951 GCCTCGGGCT CCTGAGTAGC TGGGACAACA GGTGTGCACC ACCACACCTG
7001 GCTAAATTTT TTTTTTTTCT TTGTAGAGAT AGACTCTCAC TATGTTACCT
7051 AGGCTGGTCT TGAATTCCTG GGCTCATGTA ATCCTCCTGC TGCCTTGAAC

```

7101	TCCCAAAGTG	CTGGGACTAT	AGGCATAAGA	CATCATGCCC	GGTCGGGCAC
7151	AGTGGCTCAT	GCCTGTAATC	TCAGGACTTT	GGGAGGCCGA	GACGGGCGGA
7201	TCACCTGAGG	TCGGGAGTTC	GAGACCAGCC	TGCCCAACAT	GGGAGAAACC
7251	CATCTCTACT	AAAAAAAAAA	ATACAAATTT	AGCCGAGCGT	GGTGGCACAT
7301	GCCTGTAATC	CCAGCTACTA	GGGAGGCTGA	GGCAGGAGAA	TCGCTTGAAC
7351	CCGGGAGGCT	TAGGTTGCCG	TGAGTCTGAG	TGTCACCAAT	GCACCTCAGC
7401	CTGGGCAACA	AGAGCGAAAT	TCCATCTCAA	AAAAAAAAAA	AAGAAAAAAA
7451	AGAAAAAAGA	CACCATGCCC	TATAAGTAAA	CTAGAATTAA	GGTGACTCCT
7501	AAGGAAATAA	ATAGTTTTTA	ACTGTACGAA	CTTTTGAAG	AATGGGGCCA
7551	ATTCTTTAAT	TAAATGCCAG	CTCCCTGTTT	GTGGAGAAAT	AAAAATTTTT
7601	CTTAACCCCTA	TGCCCCATT	TCCTTTCTCT	TTTATTGAAT	ATTTTTTAGT
7651	TTTAACCTATA	GTAAATACAC	CATAACGTTT	ACCATCTTAA	CCATTTTTAG
7701	GTATACAGTA	CAGTAGTGTT	CAGTACATTC	ATACTGTTAT	GCAATCAGT
7751	TCCAGAAATC	TTCATGTTTC	AAAGCTGAAA	CTCTATACCC	ATTAAACAAC
7801	TGCGTGTTCT	TCCCTCTCTC	AACCCCTGGC	AATCACCTTT	TTTTTTTTGA
7851	GACGAAGTCC	CACCTCTGCA	CCCAGGCTAG	ATGCGGTTG	CTCGATCTCG
7901	GCTCACTGCA	AGCTCCGCCT	CCCGGGTTCA	TGCCATTCTC	CTGCCTCAGT
7951	CTCCCAGACA	GCTGGGAGTA	CAGGTGCCCC	TCACACGCCG	TGCTTAATTT
8001	TTTGTATTTT	GCTAGAGAT	GGAGTTTAT	CGTGTTAGCC	AGGCTGATCT
8051	CAAACCTCTG	GCCTCAAGTG	ATCCACCCTC	CTCGGCCTCC	CAAGTGCTG
8101	GGACTACAGG	CGTGAGCCAC	TGTGCCCTGC	CAGGAAGTAG	ACTCTTGATA
8151	TTAGTTCTCT	CTGGTTGAAA	TGTTTTTAAA	AATGAAAGAG	GAGACTAAT
8201	AACAAAAAAT	CAGAAGTATA	TAAAGATTGA	TGAAGATGTG	AGATCTTTGA
8251	AACCCATGTA	TACCATTGGT	GGGAATGTGA	AACGACGCAG	CCCTGTGGAA
8301	AATGGTACAG	CAGTTACCTG	AGGTGAGGAG	TTTGAAACCA	ACCTGGCCAA
8351	CATGCAGAAA	CCCCGTCTCC	ATTAATGTGA	CAAAAATTAG	CCAGNNNNNN
8401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN		



10651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNCTCT	TTCTTCCTTT	CTTCCTTCCT
12651	CCCTTCCTCC	CTTCCTCCCT	TCCTGTCTTC	CTCTCTTTCT	CTCTTTCTTT
12701	CTTGACAGG	TCTCTCTTTG	TCTCCAGGC	TGGGGTGCAG	TGGTACAAGC
12751	ATAGCTCACA	GCAGCCTTGA	ACTCCTAGGC	TCAAGTGATC	CTCCCACGTC
12801	AGCCTCCTGA	GCAGCTGGGA	CAACGGGCTC	ATACCACCAT	GCCTGGCTAA
12851	TTTTTTAATT	TTTCGTAGAG	ACAAGTCTTT	GTTATATTGC	CCAGGCTGGT
12901	CTCAAACCTC	TGGGCTCAAA	TGCTTCTCTC	ACCTCAGCCT	CCCACGTGGC
12951	TGGGATTACA	GGCATGAGCC	ACTGCACGCC	ACTCAACACT	CCACAAATGT
13001	TGATGCCATT	ATGTTTTGTG	AACTAGTGTC	CCTGGCACCC	GAGACTTGTA
13051	CTCCACACTC	GAGGACCAAA	TAGACTGGGG	TGGGAAGGGG	TTTATAGTTT
13101	CATTATTATT	TCCCCTCAGG	GCACGGAGGT	CTTCCCCCTC	CTTGGCTCCA
13151	TCCTGCATGA	CCCCAACATC	TTCAAGCACC	CAGAAGAGTT	CAACCCAGAC
13201	CGTTTCCTGG	ATGCAGATGG	ACGGTTCAGG	AAGCATGAGG	CGTTCCTGCC
13251	CTTCTCCTTA	GGTATCTGCT	GCAGCCCTGG	GTATCACAAG	CAGGTGCTGG
13301	CGAACTCCAG	GCATCTGTGC	CAGCTGGGGG	CACCCTTCTG	CACCTGCGGC
13351	TTACTGTTGG	CTCCTCCACC	TGCTGTTCCC	CCCGTGGGCC	TGGGTGTGAG
13401	GAATACTGAC	TCAGCCCTCT	CTCTCTCTCT	CTCCTCACCA	GGGAAGCGTG
13451	TCTGCCCTGG	AGAGGGCCTG	GCAAAAGCGG	AGCTCTTCCT	CTTCTTCACC
13501	ACCATCCTAC	AAGCCTTCTC	CCTGGAGAGC	CCGTGCCCGC	CGGTACACCC
13551	TGAGCCTCAA	GCCCACCGTC	AGTGGCCTTT	TCAACATTCC	CCCAGCCTTC
13601	CAGCTGCAAG	TCCGTCCCAC	TGACCTTCAC	TCCACCACGC	AGACCAGATG
13651	AAGGAAGGCA	ACTTGGAAGT	GGTGGGTGCC	CAGGACGGTG	CCTCCAGCCT
13701	CAACAGTGGG	CATGGACAGG	GTTAATGTCT	CCAGAGTGA	CACTGCAGGC
13751	AGCCACATTT	ACACGCCTGC	AGTTGTTTTT	CGGAGTCTGT	CCCACGGCCC
13801	ACACGCTCAC	TTGACTCATG	CTGCTAAGAT	GCACAACCGC	ACACCCATAC
13851	ACAACCTACAA	GGGCCACAAA	GCAACTGCTG	GGTTAGCTTT	CCACAGACAT
13901	AAATATAGTC	CATCTGCAAT	CACAAGCACA	TAGCCAGGTA	ACCCCAAC
13951	TCCCCCTGGAT	CTGCAGCCCA	CACGTGGGAG	TCTGGCTGTC	ACCTTCACAA
14001	GCCACAGAAA	CGGCCACACA	TGTTACACAG	TCACACGCCC	TCTCCATTCA
14051	TGGAACCTCT	CAGTGTCCCT	GTCCCTGGTG	CCTGGCACAG	GGAACAGCAT
14101	GCCCCCTCCG	GGGTCATGCC	ACCCAGAGAC	TGTCGCTGTC	TATGGCCCCA
14151	ACTCATGCTC	CCTCTCTTGG	CTACACCACT	CTCCCAGCCT	GTGACCACCG

```

14201 ATGTCCACAC ACCCCCAACC ACTTGTCCAC ACAGCTACCC ACGTACGACA
14251 TCGTCCTGGC TCCCCAGAGT ATCTTCCCAC TGAGACACGC CGCCCCCACA
14301 GAGGCACAGT CCCCAGCCAC CTCTGCAACT GCAGCCCTCA GTCACCCCTT
14351 TTTAAGCACC CTGATTCTAC CAAATGCAAA CACATCTGGG TCTGCGATTA
14401 TGCACAGAGA CTTTGGACAT ACGAGGACCC TCAGACCGGA GGAACACCTG
14451 CCCAACCCCA ACACGTGCTT ATGTAACCAC GTGGAAAGCG GCCCTGCTG
14501 CCCCTCCACA CACACATACA CACTCACTGA TCTACAGCCC CTGTTGCGCG
14551 TCAGAGTCCC CACTAGACCC AGTGGGAAGG GTTAGAGACC AAGTAGGGGC
14601 CAGTTTCCAA TTCACCCGTG CAGGGAGTGA GCCGGATCTG ACGTTCCTTG
14651 TGACTTAAGG GTCCGGCTTG GGAATTAAAG TTTGTTTCTG GCCTTTAGCC
14701 TACTGCGTGT GTGACCCGTG TCAGTCACTG TGAGTAAGGG GTGGGGACAG
14751 GGGAGTCCAC CCCTCCCTG AGGCTGGGCG GGAGCTGAAA AACATGGCCA
14801 CCGCCCAACC TGCTGTGTA CATCAGGACC AGATGTGGAG CTGGGAGGAG
14851 GGGCAGGGCT GGTGACGCCC TGGGCTCAT TTCCAAAAG GGCCAAGGTG
14901 TCCGCGCGTG GGAAGTGGGC AAGGAGGGGG TAACCCAAGC TGGACTGTGG
14951 ACCTTGGGGG CTTCTCAGC CAGGGAGAGC CTGAAGCCAA CTAGATCCAG
15001 ACCCTAGAGA CTCTTCAAAC TTGAGTACAG GAACTAGCTT GCAACACAGA
15051 CTCTAAGCCC ACTCCCATT CTTCCACCCT TTTTCTCTTG CCTCCCTTC
15101 AGCTAATTTT TGTATTTTGA CAGAGGCATT TGTAATTTT CTTTCTTTT TTTTTTTTT
15151 TTTTTTTTGA GACGGAGTCT CACTCTGTCA CCCAGGCTGG AGTGCAGTGG
15201 TGTGATCTTG GCTCACTGCA GCCTCCGCTT CCGGTTCAA GCCATTCTCC
15251 TGCCTCAGCC TCCCAAGTAG CTGGGATTAC AGGTGTGTGC CACCACGCCC
15301 AGCTAATTTT TGTATTTTGA GTAGAGATGG GGTTCACCA TGTGCGCCAG
15351 GCTGGTCTCG AACTCCTGAC CTCAGATGAT CTGCCAGTCT CGGCCTCCCA
15401 AAATGCTGGG ATTACAGGCG TGAGTCGCTA CTAGATAAAT TTCTTATCTA
15451 GCAAAGAAGT TTGCAACAT ACGCAAAAGT AGAAAGATAC AATGAGCCCC
15501 CAGGTGCCCC TCACCCAGCC TCATTTCAT AGTCATCAAC TTTCTGCAGC
15551 TTTTACTTCA TCTATATCCT TTTCTGCCTC TTTTTTTTTT TTTTATTTT
15601 GAGATAGGGT TTTGCTTTGT TGCCCAAGCT GGGGTGCAGT AGCATGATCT
15651 CATAGTTCAC TGGGGCTTCA GACTCCTAGG CTCAAGTAT CCTCCCGCTT
15701 CGGCCTCCAA GCAGCTGGGA CTACAGATGC GTGCCACCAC ACCCAGCTAA
15751 ATTTCTTATT TTTATTTTCT ATAGAGAAAG TCTCACTATA CAGCCCTGTG
15801 CTGGTCTCAA ATTCCAGGCC TCAAGAGTTT CCATCCAGC CTCCCAAAGT
15851 GCTGGGATTA TAGGCGTGAG TCACTGCACC CTGCCCTAAT ATTTTATTT
15901 TATCTATTGC TTTTATTTA CTTATTTATC TTTTATTTT GAGACAGAGT
15951 CTCACCTGTG GGGCCATGCT GGAGTGCAGT GGCATCATCT CGGCTCACTG
16001 TAACCTCCGC CTCTTAGGTT CAAGCAGTTC TCCTGCCTTG ACCTCCCGAG
16051 TAGCTGGAAT TACAGGTGCC TGCCACCAAG CCTGGCTAAT TTTTTATTT
16101 GTAGTAGAGA TGGGGTTTGT CCATGTTGAC CAGGCTGGTC TCGAACTCCT
16151 GACCTCAGGT GATCTGCCCA CCTTGGCCTC CCAAAGTGCT GAGATTACTG
16201 GTATGAGCCA CCGTGCCCTG CCACCTATTG CTTTTTAAAG ATTATTTTTT
16251 TATTATTATT ATTTTTTTAT TTGCAGATGG AGTTTCGCTA TTGTTGCCCA
16301 GGCTGGAGTG CAATGGCGTG ATCTCAGCTC ACCGCAACCT CCGCCTCCCA
16351 GGTTCAGCG ATTCTCCTGC CTCAGCATCC CTAGTAGCTG GGATTACAGG
16401 CATGACCAC CATGTCCAGC TAATTTTGTA TTTTATAGT AGACGAGGT
16451 TCTCCAGGTT GGTGAGGCTG GTCTCAAAC CCACACCTCA GGTGATCCGC
16501 CCATCTCGGC CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACCGCGCC
16551 TGGCCTTAA GATTATTTTA AGGCAAATTA CAGAAAGCAA TTTAATGCAC
16601 ATTTCTGAGA GTTAAAGATA TTTTGGCCCT TGACATTTTA TGAGGACAGT
16651 TTTCAAACAT GCAGCAAAGT TGAGGGAATT GTACAAGGAA CACCTGTGTC
16701 ACTCCTGCCT CAGTCTCCCA AGCAGCTGGG ACTACAGGTG CCCGTCACCA
16751 CGCCTGGCTA ATTTTTTGTA TTTTATAGT AGATGGAGTT TCATCGTGTT
16801 AGCCAGGCTG ATCTCAAAC CCTGGCCTCA AGTGATCCAC CCGCTCGGC
16851 CTCCCAAGTG CTGGGACTAC AGGCGTGAGC CACTGTGCCT GGCCAGGTAA
16901 GTAGACTCTT GATATTAGTT CTCTCTGGTT GAAATGTTTT TAAAAATGAA
16951 AGAGAATGAC TAATAACAAA AACACAGAAA GTTATAAGGA TTGATGAAGA
17001 TGTGGAGACT TTGAAACCCA TGTATACCAT TGGTGGGAAT GTGAAACGAC
17051 GCAGCCCTGT GGAATAATGGT ACAGCAGTTA CCTGAGGTCA GGAGTTTGAA
17101 ACCAACCTGG CCAACATGCA GAAACCCCGT CTCCATTAAA TGTACAAAAA
17151 TTAGCCAGGC ATGGTGGTGC GCACCTGTAA TCCAGCTAC TCGGGAGGCT
17201 GAGGCAGGAG AATTGCTTGA ACCCAGGAGG CGGAGGTTGC AGTGAGCCGA
17251 GATCGTGCCA CTGCACTCAG CCTGGGCAAC AAAGCAAGAC TCTGTCTCAA
17301 AAAAAAAAAG TCTACTTCCC AACCTTCCCA AAAATTTATC TAAACCCCGT
17351 GACAAAACCT TAACCTGTGT TTCCGACCCC AGGCTTGGCT GTTCTGGACA
17401 TTTACTTCCC AAAGGCTGTG TTCTCTCAGC CCTCTGCCT GGTTCCTTTC
17451 AGGAGGAACA AAACCCAGGC ACAGAATTCA CCAACAAGAA CATGCTGATG
17501 ACAGTCATTT ATTTGCTGTT TGCTGGGACG ATGACGGTCA GCACCACGGT
17551 CGGCTATACC CTCTGTCTCC TGATGAAATA CCTCATGTG CAAAGTAAGA
17601 GCCTTTTCCA CTTGCCAGGC CTTGGGAACA GAAGTCAGGG TTCTAGGCTG
17651 AGCAAGGTGG CTCACGCCTA TAATCCAGC ACTTTGGGAG GCTGAGGCGG
17701 GCTGATCACT TGAGAATAGG AGTTTAAGAC CAGCCGCGCA ACACAGTGAA

```

FIGURE 3, page 5 of 10

17751 AC

FEATURES:

Start:	1999	6349
Exon:	1999	2167
Intron:	2168	5236
Exon:	5237	5386
Intron:	5387	5919
Exon:	5920	6080
Intron:	6081	6166
Exon:	6167	6349
Intron:	6354	13542
Exon:	6354	6636
Intron:	6637	13087
Exon:	13088	13261
Intron:	13262	13441
Exon:	13442	13542
Intron:	13547	13648
Exon:	13547	13648
Stop:	13649	

SNPs:

DNA Position	Major	Minor
2226	G	C A
2226	C	A G T
3081	A	T G C
3788	A	T
3979	T	C G
5056	T	C G
5213	T	G A C
5508	A	C T G
5857	C	G A T
6385	C	T G
6813	C	T A
7853	C	T
12973	G	C T
12973	A	G T
13012	G	A C T
13072	A	T C G
13370	C	G
13682	C	T A
14631	C	T G

Context:

DNA  
Position

2226	TGACAGGGGCCATGATGGAGACACCTTGGATCGAAGAGGTCACAGCACCCCTCCTCTTTCT TCCTCCCTACCCCACTGAGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACC CTGGCGGCCCTGTGGTGGTCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTCA GGCTGAGGAGTTTACGCGGCCGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCA TGGAAGTCAAGGGCTGCTAGGCCCTCCGCTCACAGCCTGCCACCACTTACTGGTGTGTG [G, C, A] CCTTTGCACATGGCTTAGTCCCTCTGTTGCCTCATCTGTCAAATGGAGTGATAACAGTGC CCATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGG GTGGATCACTTGAGGTCAGGAGTTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCTGTCT CTACTAAAAATATAAAATTAGCTGGGCATGGTGGTGCCTACCTGTAATCCCAGATACTT GGGAGGTTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGA
2226	TGACAGGGGCCATGATGGAGACACCTTGGATCGAAGAGGTCACAGCACCCCTCCTCTTTCT TCCTCCCTACCCCACTGAGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACC CTGGCGGCCCTGTGGTGGTCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTCA GGCTGAGGAGTTTACGCGGCCGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCA TGGAAGTCAAGGGCTGCTAGGCCCTCCGCTCACAGCCTGCCACCACTTACTGGTGTGTG [C, A, GsT]

CCTTTGCACATGGCTTAGTCCCTCTGTTGCCTCATCTGTCAAATGGAGTGATAACAGTGC  
CCATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGG  
GTGGATCAGTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCT  
CTACTAAAAATATAAAAATTAGCTGGGCATGGTGGTGCCTACCTGTAATCCAGATACTT  
GGGAGTTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGA

3081 AAACAGTTCTCTGTATTACAACAGAAAGCAGGAGGCCATGCTGGGTGCTGCCAGGAAC  
CAGTAGTAACATAAGACAGCACCGGTGCTGCTTCCCCAGCGCACCTAGGCCAGTGGGGAAA  
CAGACTCACCACACAGTCCAGCCCAGAGTGGTCAGGGCCAAGATGGGGAAGCACGGGGA  
GAAAGGTCAGGGTGGGATGGGAGGGGTGAGGCAAGAGGGGTGAGGGCCAGGCTGAGGG  
AAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGACCCGGCATGTTGGTGAGGGAG  
[A, T, G, C]  
TTCAGGAAGTCTTCTGGAAGAGAGGCTGTGCGAGCTGAGACTCATAAGATGAGTGGGGA  
GGGTGTTCCAGGCAGAAAGACAGCACCTACAAAAGCATGACTTTGAGAGAAGCATTTCAT  
CCATTCACTGATGAATTTTCAGACTGGGCACGCTGGCTCATGCCTGTAATCCAGCACT  
TTGGAAGGCTGAATGGGAGGATGACTTGAAGCTAGGCATTGTGACAAGCCTGGGCAAC  
ATGGTGAGACCTGCTCCACAAAACAAACAAACAAACAAAAATCATTATACCTGGTAC

3788 ATCTTAAATGAGAGGTACAGATTGAGAGCAAACACAGGCACAGGCATATGTACGAGG  
GTAAAGAGGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCCGGGACATGAAG  
GATGAACGAGTTAGTTACCAAGGATGGGATGGAAGGGGTGAGAGTGATGGAGGCAGAG  
GGAACGTCAGGATCATAGGCCTAGACAGGGGATCCTGACGCCCTTGAGGAAGTGAGAGAA  
GACCAGCGCAGTCGTAGTGGGTTAAGTAACAAGCTGAGAAGCCAGGGAATCCCTGGTC  
[A, T]  
TGCAGGGCCTGTGAGTCAGTCAGAGTGTGTTGGGCTTTTGTGTTTCTGGGAGCAGTCGA  
TTTTAAGCAGGGAACAGCTGTATTTCAGAGTTGGGAAGATCCTGTGGTGTCTGCTGAAGG  
GGATGAAACTGGAGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGATGGGGAGGC  
TGGGAGGTCCCGGCTGATGGACAGGGCTGGGGCCAGGGGATGGGGAGGAAGGAGTAATT  
GGGAGAGCCTGGGCTCTGCCCAGGAATGGATGGTGGGCTGAACAGGGAGAGGAGAG

3979 ATCATAGGCCTAGACAGGGGATCCTGACGCCCTTGAGGAAGTGAGAGAAGACCAGCGAG  
TCGTAGTGGGTTAAGTAACAAGCTGAGAAGCCAGGGAATCCCTGGTCATGCAGGGCCT  
GTGAGTCACGTCAGAGTGTGTTGGGCTTTTGTGTTTCTGGGAGCAGTCGATTTTAAGCAG  
GGAACAGCTGTATTCAGAGTTGGGAAGATCCTGTGGTGTCTGCTGAAGGGGATGAACT  
GGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGATGGGGAGGCTGGGAGGTCC  
[T, C, G]  
CGGTGATGGACAGGGCTGGGGCCAGGGGATGGGGAGGAAGGAGTAATTGGGAGAGGCCT  
GGGGCTCTGGCCGAGGAATGGATGGTGGGCTGAAACAGGGAGAGGAGATGCTTAGGCC  
ACTTTGGAACACAGTAGGGCAAGGACAGGAGACACCCAGGGGAAGTGCCCAAGAGACCA  
CGACAGGCTGGCATTGGACAGGGAAGGTCTGTCTGGAGCAGGTGTCTTGGATAAGGGAGG  
AAAATGGTGAGTTCATCCTCCTCCCTCTCTGTTCACCTCTAAACTACATGGGGCACA

5056 AGTTTGATTATGCATTTATTGAGCACCTACTGAGTCCTCATCCCTGGGCTAGGCTGGAAT  
GGAATCAGATGGAGCCTGAAGAGTCCCCCTCAGGGAACCTCACTAGAAAGAAGGAGGAAT  
CGGCCGGGCGCGGTGGCTCACGCTGTAAATCCCAACACTTTGGGAGGCTGAGGTGGGTGG  
ATCACAAGGTCAGGAGATCGAGACCCTCTGGCTAACACAGTGAACCCCATCTCTACTA  
AAAATACAAAAAATGAGCCAGGCATGTTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG  
[T, C, G]  
CTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCAGAGGTTGCACTGAGACGAGATCACGC  
CACTGCACCTCCAGCCTGGGCAACAGAGCGAGATTCCGTCTCAAAAAAAAAAAGAAAGAAA  
GGAAGAAGGGGGAATGGGGGAGAGGGGCGGTCCCTTTTGGAGTCTAGCCTTCTCGCGCAG  
GGGTTTCTTCTTCAACGGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGCTC  
TGCGGGACCTGGGCATGGGAAGCGAGAAGGCCAGGAGCTGATCCAGGCAGGCCCCGT

5213 CTTTGGGAGGCTGAGGTGGTGGATCACAAGGTCAGGAGATCGAGACCATCCTGGCTAAC  
ACAGTGAAACCCATCTCTACTAAAAATACAAAAATGAGCCAGGCATGGTGGCGGGCGC  
CTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCAGA  
GGTTGCACTGAGACGAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGCGAGATTCC  
GTCTCAAAAAAAAAAAGAAAGGAAGAAGGGGAATGGGGAGAGGGGCCGGTCCCT  
[T, G, A, C]  
TTTGAAGTCTAGCCTTCTGCGCAGGGGTTTCTTCTTCAACGGGGAGCGGTGGAGGCAGCT  
GAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATGGGGAAGCGAGAAGCGAGGA  
GCTGATCCAGGCGGAGGCCCGGTGTCTGGTGGAGACATTCCAGGGGACAGAAGGTGAGCA  
TGGCGGGGTCAACCCAGGCTCTCCAGCCGAGTGAAAGGGAAACTCTCTACTGTGGCTG  
GGGGTGGCCCCAACCCAGGCTCTGGAATGGGCAGGAGGGGAAGCCTTGAAGTCTAGGGCT

5508 TCCCTTTTGTAGTCTAGCCTTCTGCGCAGGGGTTTCTTCTTCAACGGGGAGCGGTGGAG  
GCAGCTGAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATGGGGAAGCGAGAAGG  
CGAGGAGCTGATCCAGGCGGAGGCCCGGTGTCTGGTGGAGACATTCCAGGGGACAGAAGG  
TCAGCATGGCGGGGTCAACCCAGGCTCTCCAGCCGAGTGAAAGGGAAACTCTCTACTG  
TGGCTGGGGGTGGCCCCAACCCAGGCTCTGGAATGGGCAGGAGGGGAAGCCTTGAAGTCT

FIGURE 3, page 7 of 10

[A, C, T, G]  
GGGCTGGCCTGGGGGTTCTGTTCACTGCCACCTTCTGTCTCTGTCCCACCTGTCTCTCCGA  
GGCTGTGCATGACATCTCTGTGTCTCTGGTGCTATCATCCCATTTCTCCTGGGTCTC  
CATCTCTCTCTGTCTCTTTTCTTTCTCTCTCCTTTCTCTATTTTTGGGCCCTCAGT  
CTATCTCTGTTTCTGTCTCCCTGTCTGTGTGATGGTCACTCTGTTTCTTCTCCTGTCT  
GTTTCTCTGTCCCTATCTGTCTGTATCCTTCTTGGCTGTTTAGCTCTCTCCCTGCGCTG

5857 CTGTCTCTCCGAGGCTGTCATGACATCTCTGTGTGTCTCTGGTGCTATCATCCCATTC  
TTCCTGGGTCTCCATCTCTCTCTGTCTCTTTTCTTTCTCTCTCCTTTCTCTATTTTT  
TGGGCCCTCAGTCTATCTCTGTTTCTGTCTCCCTGTCTGTGTGATGGTCACTCTGTTTCT  
TTCTCCCTGTCTGTTTCTGTCTCCCTATCTGTCTGTATCCTTCTTGGCTGTTTAGCTCT  
CTCCCTGCGCTGTCCATCCATCTTTCCCTGCTCCTGTCTCTCTGTTGGGTTGAGC  
[C, G, AsT]  
CCAACCTGCTCCCTCTGCTGGCTCCATCACAGCCTACCTCCCTGCCCCATTTCCCCC  
AGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAACGTAGTCTGCTC  
CCTCCTCTTTGGCCTCCGCTTCTCCTATGAGGATAAGGAGTTCCAGGCCGTGGTCCGGGC  
AGCTGGTGGTACCTGCTGGGAGTCAGCTCCAGGGGGGTGAGTGAGTGGGTGGGACCC  
CTCTCCAACCTACCTTCCCTGAAGGTTCTGCCAAGGTCCCATGAGAACTAGCTGCCCTTC

6385 GTGGGTGGGACCCCTCTCCAACCTACCTTCCCTGAAGGTTCTGCCAAGGTCCCATGAGAA  
CTAGCTGCCCTTCTCCCCACAGACCTACGAGATGTTCTCCTGGTTCTGCGGCCCTGCC  
AGGCCCCACAAAGCAGCTCCTCCACCACGTGACACCTTGGCTGCCCTTCCAGTCCGGCA  
GGTGACAGCAGCACAGGGGAACCTGGATGCTTCGGGCCCCGACGTGACCTTGTGATGC  
CTTCTGCTGAAGATGGCACAGGTGTGGGAAGGTCAGGGACCCCTCTCTGAATGGGC  
[C, T, G]  
TGGTGACCTGGCAGGTCCAGCCAGGTGTCCCTGGGGACCTCAATTGGGTTCTCTCTCT  
TTCTCTCTCTGTCATGTCTGTGTGATGATGATGTCTCTGTGCATGTGTGTGTCATCCCTTC  
TCTGCACATCTGTGTGGCCTTTCAGGGCGTGTCTCTCACTGCCTCTCCCGCCCCGAC  
CTGGGCATTTGTGCGGGGCTGTCTGTCTCTCCAGCATCTCTCCTTTTCTCCCTCCCAAC  
TCGGCCCTTGTGTTTCCAGGCCCATGCCAGGGTCTTACACCAGCAATCCCAGGATCACT

6813 ATCTGTGCTGGCCCTTTCAGGGCGTGTCTCTCACTGCCTCTCCCGCCCCGACCTGGGCA  
TTTGTGCCGGGCTGTCTGTCTCTCCAGCATCTCTCCTTTCTCCTCCCACTCGGGCC  
TTGTGTTTCCAGGCCCATGCCAGGGTCTTACACCAGCAATCCCAGGATCACTTCATCCC  
ATCCCCTGACGCTCCCCAGACTTTTATGTAAATTCACAATTTTATGTGAATTATGGTCA  
TTTATTAGGAAGCCTTGCAATATCAAGTTATGTTAATAAAGTCCACTTTATTATATAT  
[C, T, A]  
AGAACAATATTTCTTTCTTTTCTTTTCTTTTCTTTTAAAGAGACAGGATCTCTTTC  
TGTTCGCCAGGCTAGAGTACAGTTGCAAAATCATAGCTCACTGCAACCTTGAATCCTGG  
GCTCAAGCAATCCTCCTGCTCGGGCTCCTGAGTAGCTGGGACAAAGGTGTGCACCACC  
ACACCTGGCTAAATTTTTTTTTTTCTTTGTAGAGATAGACTCTCACTATGTTACCCAGG  
CTGGTCTTGAATTCCTGGGCTCATGTAATCCTCCTGCTGCCTTGAACCTCCCAAGTGTG

7853 TCTTTAATTAATGCAGCCTCCCTGTTTGTGGAGAAAGAAAATTTTCTTAACCTATT  
GCCCCATTTCTTTCTCTTTTATGAATATTTTATAGTTTAACTATAGTAAATACACA  
TAACGTTTACCATCTTAACCATTTTATAGGTATACAGTACAGTAGTGTTCAGTACATTCA  
ACTGTTATGAATCAGTCTCCAGAACTCTTATGTTGCAAAAGCTGAAACTCTATACCCAT  
TAAACAACCTGCTTCTCCTCCTCCAACCCCTGGCAATCACCTTTTTTTTTTTGAGA  
[C, T]  
GAAGTCTCACTCTGTCAACCAGGCTAGAGTGCAGGTGGCTCGATCTCGGCTCACTGCAAGC  
TCCGCTCCCGGGTTCATGCCATTCTCCTGCCTCAGTCTCCCAAGCAGCTGGGACTACAG  
GTGCCCGTCAACACGCTGGCTAATTTTTTTGTATTTTATAGTAGAGATGGAGTTTCATCGT  
GTTAGCCAGGCTGATCTCAAACCTCCTGGCCTCAAGTATCCACCGCTCGGCTCCCAA  
AGTGCTGGGACTACAGGCTGAGCCACTGTGCTGGCCAGGAAGTAGACTCTTGATATTA

12973 CTGTCTCTCTCTCTTCTCTCTTCTTTCTTTGACAGGGTCTCTCTTTGTCTCCAGGCTG  
GGGTGCAGTGGTACAAGCATAGCTCACAGCAGCCTTGAACCTCCTAGGCTCAAGTGATCCT  
CCCACGTTACGCTCCTGAGCAGCTGGGACAACGGGCTCATACCACCATGCTGGCTAATT  
TTTTAATTTTTCGTAGAGACAAGGTCTTGTTATATTGCCAGGCTGGTCTCAAACCTCCTG  
GGCTCAAATGCTTCTCTACCTCAGCCTCCCAAGTGGCTGGGATTACAGGCATGAGCCAC  
[G, C, T]  
GCACGCCACTCAACACTCCACAAATGTTGATGCCATTATGTTTTGTGAAGTGTGCTCCT  
GGCACCAGAGACTTGTACTCCACACTCGAGGACCAATAGACTGGGGTGGGAAGGGGTTT  
ATAGTTTCATTATTATTTCCCTCAGGGCACGGAGTCTTCCCCCTCCTTGGCTCCATCC  
TGCATGACCCCAACATCTTCAAGCACCAGAGAGTTCAACCCAGACCGTTTCTGGATG  
CAGATGGACGTTTCCAGGAAGCATGAGGCGTTCTGCCCTTCTCCTTAGGTATCTGTGCA

12973 CTGTCTTCTCTCTTTCTCTCTTTCTTTCTTTGACAGGGTCTCTCTTTGTCTCCAGGCTG  
GGGTGACGTGGTACAAGCATAGCTCACAGCAGCCTTGAACCTCCTAGGCTCAAGTGATCCT  
CCACAGTCAGCCTCCTGAGCAGCTGGGACAACGGGCTCATACCACCATGCGCTGGCTAATT  
TTTTAATTTTTCGTAGAGACAAGGTCTTGTATATTGCCCAGGCTGGTCTCAAACCTCCTG  
GGCTCAAATGCTTCTCTACCTCAGCCTCCACGCTGGCTGGGATTACAGGCATGAGCCAC  
[A, G, T]  
GCACGCCACTCAACACTCCACAAATGTTGATGCCATTATGTTTTGTGAAGTAGTGCCCT  
GGCAGCCGAGACTTGTACTCCACACTCGAGGACCAATAGACTGGGGTGGGAAGGGGTTT  
ATAGTTTCATTATTATTTCCCTCAGGGCAGGAGGTCTTCCCCCTCCTTGGCTCCATCC  
TGCATGACCCCAACATCTTCAAGCAGCCAGAGAGTTCAACCCAGACCGTTTCTGGATG  
CAGATGGACGGTTCAGGAAGCATGAGCGTTCTGCCCTTCTCTTAGGTATCTGCTGCA

13012 CTCTCTTTGTCTCCAGGCTGGGGTGCAGTGGTACAAGCATAGCTCACAGCAGCCTTGAA  
CTCTTAGGCTCAAGTGATCCTCCACGTCAGCCTCCTGAGCAGCTGGGACAACGGGCTCA  
TACCACCATGCGCTGGCTAATTTTTTAATTTTTCGTAGAGACAAGGTCTTGTATATTGCC  
CAGGCTGGTCTCAAACCTCCTGGGCTCAAATGCTTCTCTACCTCAGCCTCCACGCTGGCT  
GGGATTACAGGCATGAGCCACTGCACGCCACTCAACACTCCACAAATGTTGATGCCATTA  
[G, A, CsT]  
GTTTTGTGAAGTAGTGCCCTGGCAGCCGAGACTTGTACTCCACACTCGAGGACCAATA  
GACTGGGGTGGGAAGGGGTTTATAGTTTCATTATTATTTCCCTCAGGGCAGGAGGTCT  
TCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCAGCCAGAGAGTTCA  
ACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCTGCCCT  
TCTCCTTAGGTATCTGCTGCAGCCTGGGTATCACAAGCAGGTGCTGGCGAACTCCAGGC

13072 CTCCTAGGCTCAAGTGATCCTCCACGTCAGCCTCCTGAGCAGCTGGGACAACGGGCTCA  
TACCACCATGCGCTGGCTAATTTTTTAATTTTTCGTAGAGACAAGGTCTTGTATATTGCC  
CAGGCTGGTCTCAAACCTCCTGGGCTCAAATGCTTCTCTACCTCAGCCTCCACGCTGGCT  
GGGATTACAGGCATGAGCCACTGCACGCCACTCAACACTCCACAAATGTTGATGCCATTA  
TGTTTTGTGAAGTAGTGCCCTGGCAGCCGAGACTTGTACTCCACACTCGAGGACCAAT  
[A, T, CsG]  
GACTGGGGTGGGAAGGGGTTTATAGTTTCATTATTATTTCCCTCAGGGCAGGAGGTCT  
TCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCAGCCAGAGAGTTCA  
ACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCTGCCCT  
TCTCCTTAGGTATCTGCTGCAGCCTGGGTATCACAAGCAGGTGCTGGCGAACTCCAGGC  
ATCTGTGCCAGCTGGGGGACCCCTTCTGCACCTGGGCTTACTGTTGGCTCCTCCACCTG

13370 ATAGACTGGGGTGGGAAGGGGTTTATAGTTTCATTATTATTTCCCTCAGGGCAGGAGG  
TCTTCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCAGCCAGAGAGT  
TCAACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCTGCTG  
CCTTCTCCTTAGGTATCTGCTGCAGCCTGGGTATCACAAGCAGGTGCTGGCGAACTCCA  
GGCATCTGTGCCAGCTGGGGGACCCCTTCTGCACCTGGGCTTACTGTTGGCTCCTCCAC  
[C, G]  
TGCTGTTCCCCCGTGGGCTGGGTGTGAGGAATACTGACTCAGCCCTCTCTCTCTCTCT  
CTCTCACCAGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAAGCGGAGCTCTTCTCT  
CTTCTTCAACCATCTTACAAGCCTTCTCCCTGGAGAGCCGCTGCCCGCGGTACACCC  
TGAGCCTCAAGCCACCGTCACTGGCCTTTTCAACATTTCCCCAGCCTTCCAGCTGCAAG  
TCCGTCCCACTGACCTTCACTCCACCACGACAGCAGATGAAGGAAGGCAACTTGAAGT

13682 CCGTGGGCTGGGTGTGAGGAATACTGACTCAGCCCTCTCTCTCTCTCTCTCTCACCAG  
GGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAAGCGGAGCTCTTCTCTCTTTCACCA  
CCATCTTACAAGCCTTCTCCTTGGAGAGCCCGTGCCCGCGGTACACCTGAGCCTCAAG  
CCCACCGTCACTGGCCTTTTCAACATTTCCCCAGCCTTCCAGCTGCAAGTCCGTCCCACT  
GACCTTCACTCCACCACGACAGATGAAGGAAGGCAACTTGGAAAGTGGTGGGTGCC  
[C, T, A]  
GGACGGTGCCTCCAGCCTCAACAGTGGGCTGACAGGGTTAATGTCTCCAGAGTGATACA  
CTGCAGGAGCCACATTTACACGCTGCAGTTGTTTTCCGAGTCTGTCCCACGGCCAC  
ACGCTCACTTGAATCATGCTGCTAAGATGCACAACCGCACACCCATACACAACATACAAG  
GCCACAAAGCAACTGCTGGGTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCA  
CAAGCACATAGCCAGGTAACCCACCAACTCCCTGGATCTGCAGCCACACGCTGGGAGTC

14631 GCAGCCCTCAGTCACCCCTTTTAAAGCACCCCTGATTCTACCAAATGCAAACACATCTGGG  
TCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGACCGGAGGAACACCTG  
CCCAACCCCAACACGTGCTTATGTAACCACGTGAAAGCGGCCCTGCTGCCCCCTCCACA  
CACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAGAGTCCCCACTAGACCC  
AGTGGAAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACCCTGTCAGGGAGTGA  
[C,T,G]  
CCGGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAATTAAAGTTTGTCTCTGG  
CCTTTAGCCTACTGCGTGTGTGACCCGTGTCAGTCACTGTGAGTAAGGGGTGGGGACAGG  
GGAGTCCACCCCTCCCCTGAGGCTGGGCGGGAGCTGAAAAACATGGCCACCGCCACCCCT  
GGCTGTTGACATCAGGACCAGATGTGGAGCTGGGAGGAGGGGCAGGGCTGGTGACGCCCT  
GGGCCTCATTTCCAAAAAGGGCCAAGGTGTCCGGCGGTGGGAAGTGGGCAAGGAGGGGGT

Chromosomal Map Position

ePCR to dbSTS

Site (bases)	Marker	Chr.	Organism
15155..15288	stSG46708	19	Homo sapiens

004347 228460